REMARKS

Applicant has studied the Office Action dated May 3, 2000 and has made amendments to the claims. It is submitted that the application, as amended, is in condition for allowance. Claims 1-27 are pending. Claims 1, 9, 17 and 25-27 have been amended. Reconsideration and allowance of the claims in view of the above amendments and the following remarks are respectfully requested.

The specification has been carefully amended to correct grammatical and minor typographical errors. No new matter has been added.

Claims 1, 3, 5, 8, 9, 11, 13, 16-20 and 24-27 were rejected under 35 U.S.C. §102(e) as being anticipated by Kageyama (U.S. Patent No. 5,955,693). Claims 4, 14 and 22 were rejected under 35 U.S.C. §103(a) as being unpatentable over Kageyama. These rejections are respectfully traversed.

The present invention is directed to a voice converter in which, for example, imitation of a professional singer by a karaoke player is capable of being performed. In a preferred embodiment of the present invention, the voice converter includes an extracting means for extracting a plurality of sinusoidal wave components from an input voice signal, including frequencies and/or amplitudes of the sinusoidal wave components. For example, the set of sinusoidal components may be in the form of frequency value F and amplitude value A coordinates, such as (F0, A0), (F1, A1), (F2, A2), ...(Fn, An), where n is an integer (see Figs. 2 and 3). A modulating means is provided to modulate frequencies and/or amplitudes of the sinusoidal wave components according to pitch information and/or amplitude information of a reference voice signal, as depicted in Fig. 6. The reference voice signal may, for example, represent a professional singer's voice signal that a karaoke singer is trying to imitate. After the modulation, a mixing means mixes the plurality of the sinusoidal wave components to synthesize an output voice signal having a pitch different from that of the input voice signal and influenced by that of the reference voice signal.

Accordingly, when the voice of the karaoke singer is output, the characteristics of the voice, the manner of singing, and the like, are significantly influenced by the reference voice signal. For example, referring to Fig. 6, in a system that both frequencies and amplitudes are modulated, the frequency values shown in part (4) of Fig. 6 are combined with the reference pitch information Pto and the fluctuation component Ptf to give the modulated frequency values in part (7) of Fig. 6. The modulated frequency values and the modulated amplitude values shown in parts (7) and (8) of Fig. 6 are combined by the mixing means, thereby yielding new sinusoidal components as illustrated in part (9) of Fig. 6. From these new sinusoidal components, the output voice signal is operatively formed. This allows the input voice characteristics to imitate a reference voice signal and the input voice to imitate the singing manner of a desired reference singer.

The Kageyama reference discloses a karaoke apparatus capable of changing a live singing voice to a similar voice of an original singer of a karaoke song. However, Kageyama does not disclose an extracting means for "extracting a plurality of sinusoidal wave components from [an] input voice signal, including frequencies of the sinusoidal wave components of the input voice signal" and a modulating means for "modulating frequencies of the sinusoidal wave components of the input voice signal according to pitch information" representative of a pitch of a reference voice signal (hereinafter referred to as "the frequency limitation"), as recited in amended claim 1 and similarly recited in amended claim 25. Likewise, Kageyama does not disclose an extracting means for "extracting a plurality of sinusoidal wave components from the input voice signal, including amplitudes of the sinusoidal wave components in the input voice signal" and a modulating means for "modulating the amplitude of each sinusoidal wave component extracted from the input voice signal according to the amplitude information" representative of amplitudes of sinusoidal wave components contained in a reference voice signal (hereinafter referred to as "the amplitude limitation"), as is recited in amended claims 9 and similarly recited in amended claim 26. Claim 17 incorporates both the frequency and amplitude limitations in claims 1 and 9, analyzing and modulating "a plurality of sinusoidal wave components contained in the input voice signal to derive a parameter set of an original frequency and an original amplitude, each

pair of the original frequency and the original amplitude representing a corresponding sinusoidal wave component." Amended claim 27 contains similar recitations.

Kageyama discloses a method of modifying a live singing voice to a voice similar to the original/model singer of the karaoke song. However, the voice modifying method is different from the present voice conversion using the set of sinusoidal components, such as in the form of (Fn, An). The karaoke apparatus in the Kageyama reference uses phoneme data of a model singer to modify and approximate the voice of the live karaoke singer to that of the model singer. The phoneme data represents primary characteristics of the vowels contained in the model voice of the model singer, in terms of the waveform, envelope thereof, vibrato frequency, vibrato depth and supplemental noise (see column 4, line 66 to column 5, line 2).[1] When the live singing voice is input in the karaoke apparatus in the Kageyama reference, a separating device separates the lead consonant component and the subsequent vowel component of the live singing voice. After the separation, an extracting device extracts the secondary characteristics of the subsequent vowel component, which may for example be the pitch of the separated subsequent vowel component. A substitutive vowel component is then created according to the primary characteristics of the vowels in the model voice (i.e., the phoneme data of model voice) and the secondary characteristics (e.g., the pitch of input voice). The substitutive vowel component, having the waveform of the model vowel and the pitch of the separated subsequent vowel component from the live singing voice, basically replaces the subsequent vowel component. Finally, the substitutive vowel component and is combined with the lead consonant component to synthesize an output singing voice.

In contrast, the present invention utilizes a voice conversion apparatus that extracts a plurality of sinusoidal wave components from an input voice signal, including frequencies and/or amplitudes of the sinusoidal wave components in the input voice signal. The set of sinusoidal wave components may, for example, be in the form of (Fn, An), where n is an integer. Each (Fn, An) component represents a pair of the original frequency and the original amplitude of each

---

[1] Referring to Fig. 6A of the Kageyama reference, a phrase of lyric "A KA SHI YA NO" comprises five syllables "A", "KA", "SHI", "YA" and "NO", and the phoneme data are composed of extracted vowels "a", "a", "I", "a" and "o" from the five syllables.

sinusoidal wave component. All of some of the extracted sinusoidal wave components are then modulated by reference pitch information of a reference voice signal and/or reference amplitude information of a reference voice signal. In the karaoke apparatus of Kageyama, vowel components are extracted from the input voice and then replaced with a substitute vowel component having a waveform of a reference model vowel and a pitch of the input voice.

The karaoke apparatus of Kageyama does not disclose, teach or suggest extracting a set of sinusoidal components, including frequencies of the sinusoidal wave components and/or amplitudes of the sinusoidal wave components, and modulating these components. Likewise, the karaoke apparatus of Kageyama does not disclose, teach or suggest an analyzer device that analyzes a plurality of sinusoidal wave components contained in the input voice signal to derive a parameter set of an original frequency and an original amplitude, each pair of the original frequency and the original amplitude representing a corresponding sinusoidal wave component, and a modulator device that modulates the parameter set of the sinusoidal wave components according to reference information. Therefore, it is respectfully submitted that claims 1, 9, 17 and 25-27 distinguish over the Kageyama reference. Because each dependent claim incorporates all the limitations of its base claim(s), claims depending from 1, 9 and 17 also distinguish over the Kageyama reference. Claims 3, 5 and 8 depend directly or indirectly from claim 1. Claims 11, 13 and 16 depend directly or indirectly from claim 9. Claims 18-20 and 24 depend directly or indirectly from claim 17. Rejections of claims 1, 3, 5, 8, 9, 11, 13, 16-20 and 24-27 under 35 U.S.C. §102(e) and claims 4, 14 and 22 under 35 U.S.C. §103(a) should be withdrawn.

Claims 2, 6, 10, 12 and 21 were rejected under 35 U.S.C. §103(a) as being unpatentable over Kageyama in view of Matsumoto '303 (U.S. Patent No. 5,847,303). Claims 7, 15 and 23 were rejected under 35 U.S.C. §103(a) as being unpatentable over Kageyama in view of Matsumoto '907 (U.S. Patent No. 5,963,907). These rejections are respectfully traversed.

The claimed features of the present invention are not realized even if the teachings of the Matsumoto '303 reference or Matsumoto '907 reference are incorporated into Kageyama. Matsumoto '303 is directed to a voice processing apparatus that modulates an input voice signal into an output voice signal according to a set of parameters. Matsumoto '303 discloses a voice

change parameter table of filter coefficients to control spectrum shape of varying pitch ranges for the purpose of providing more realistic sounding conversion between male and female voices (see Figs. 9 and 10; column 11, lines 3-26). An audio signal processor within the voice processing apparatus is configured by a parameter set to process the audio signal by modifying the frequency spectrum of the input voice. However, Matsumoto '303 does not disclose the inventive features of the present invention in extracting a plurality of sinusoidal wave components from an input voice signal, including frequencies of the sinusoidal wave components and modulating frequencies of the sinusoidal wave components according to pitch information representative of a pitch of a reference voice signal, as recited in amended claim 1. Likewise, Matsumoto '303 does not disclose extracting a plurality of sinusoidal wave components from the input voice signal, including amplitudes of the sinusoidal wave components and modulating amplitudes of the sinusoidal wave component extracted from the input voice signal according to the amplitude information representative of amplitudes of sinusoidal wave components contained in a reference voice signal, as is recited in amended claim 9. Claim 17 incorporates the above limitations in amended claims 1 and 8 and distinguishes over the Matsumoto '303 reference.

Matsumoto '907 is directed to a voice converter that provides pitch and formant shifting of an input voice signal. Referring to Fig. 2 of the Matsumoto '907 reference, an audio filter 325 extracts the volume level of the input voice signal, and outputs the extracted volume level as first volume data V1. A second audio filter 326 extracts the volume level of an output voice signal, and outputs the extracted volume level as second volume data V2. A difference judging circuit 322 compares the first and second volume data V1 and V2 with each other, and determines a volume gain G and a distorting factor D which is supplied to a distortion circuit 321. When the volume of the output voice after conversion is smaller than that of the input voice, the volume gain G is increased. In contrast, the subject matter of claims 7, 15 and 23 in the present invention is to change the volume of an input singing voice in matching with the variation of the volume of the voice of a model singer. This allows the volume of an output voice signal to emulate the volume variation of the reference voice signal of the model singer. Such feature is not disclosed, taught or suggested by Matsumoto '907. Additionally, Matsumoto '907 does not disclose the inventive features of the present invention in extracting a plurality of sinusoidal wave

components from an input voice signal, including frequencies of the sinusoidal wave components and modulating frequencies of the sinusoidal wave components according to pitch information representative of a pitch of a reference voice signal, as recited in amended claim 1. Likewise, Matsumoto '907 does not disclose extracting a plurality of sinusoidal wave components from the input voice signal, including amplitudes of the sinusoidal wave components and modulating amplitudes of the sinusoidal wave components extracted from the input voice signal according to the amplitude information representative of amplitudes of sinusoidal wave components contained in a reference voice signal, as is recited in amended claim 9. Claim 17 incorporates the above limitations in amended claims 1 and 8 and distinguishes over the Matsumoto '907 reference.

Applicant believes that the differences between Kageyama, Matsumoto '303, Matsumoto '907 and the present invention are clear in amended claims 1, 9 and 17, which set forth voice conversion and synthesizing apparatuses that utilize a plurality of sinusoidal wave components according to embodiments of the present invention. Therefore, claims 1, 9 and 17 distinguish over the Kageyama, Matsumoto '303 and Matsumoto '907 references. Claims depending directly or indirectly from claims 1, 9 and 17 also distinguish over the above references. Applicant further believes that the differences between Kageyama, Matsumoto '907 and the present invention are clear in claims 7, 15 and 23, which set forth apparatuses that emulate volume variation of a model singer according to embodiments of the present invention. Therefore, the rejection of claims 2, 6, 7, 10, 12, 15, 21 and 23 under 35 U.S.C. § 103(a) should be withdrawn.

In view of the foregoing, it is respectfully submitted that the application and the claims are in condition for allowance. Reexamination and reconsideration of the application, as amended, are requested.

If for any reason the Examiner finds the application other than in condition for allowance, the Examiner is invited to call the undersigned attorney at (213) 488-7100 should the Examiner believe a telephone interview would advance the prosecution of the application.

Respectfully submitted,

Date: August 31, 2000

By: _____
Roger R. Wise
Registration No. 31,204
Attorney for Applicant

PILLSBURY MADISON & SUTRO LLP
725 South Figueroa Street, Suite 1200
Los Angeles, California 90017-5443
Telephone: (213) 488-7100
Facsimile: (213) 629-1033